PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

C12N15/22

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 4 :
C07K 13/00, 15/00, A61K 37/00
C12P 21/00, 21/02, C12N 15/00
C07H 15/12

A1

(11) International Publication Number: WO 88/ 00205

(43) International Publication Date: 14 January 1988 (14.01.88)

(21) International Application Number: PCT/US87/01537

(22) International Filing Date: 30 June 1987 (30.06.87)

(31) Priority Application Numbers: 880,776
943,332
028,285
031,346

(32) Priority Dates: 1 July 1986 (01.07.86)
17 December 1986 (17.12.86)
20 March 1987 (20.03.87)
26 March 1987 (26.03.87)

(33) Priority Country: US

(71) Applicant: GENETICS INSTITUTE, INC. [US/US]; 87 CambridgePark Drive, Cambridge, MA 02140 (US).

(72) Inventors: WANG, Elizabeth, A. ; 136 Wolf Rock Road, Carlisle, MA 01741 (US). WOZNEY, John, M. ; 59 Old Bolton Road, Hudson, MA 01749 (US). ROSEN, Vicki, A. ; 344 Marlborough Street, Apartment 4, Boston, MA 02116 (US).

(74) Agents: BAK, Mary, E. et al.; Genetics Institute, Inc., 87 CambridgePark Drive, Cambridge, MA 02140 (US).

(81) Designated States: AT (European patent), AU, BE (European patent), BG, BJ (OAPI patent), CF (OAPI patent), CG (OAPI patent), CH (European patent), CM (OAPI patent), DE (European patent), DK, FI, FR (European patent), GA (OAPI patent), GB, GB (European patent), HU, IT (European patent), JP, KR, LU (European patent), ML (OAPI patent), MR (OAPI patent), NL (European patent), NO, SE (European patent), SN (OAPI patent), SU, TD (OAPI patent), TG (OAPI patent).

Published
With international search report.
With amended claims.

(54) Title: NOVEL OSTEOINDUCTIVE COMPOSITIONS

(57) Abstract

Human and bovine bone inductive factor products and processes. The factors may be produced by recombinant techniques and are useful in the research and treatment of bone and periodontal defects.

NOVEL OSTEOINDUCTIVE COMPOSITIONS

The present invention relates to novel proteins and processes for obtaining them. These proteins are capable of inducing cartilage and bone formation.

Background

Bone is a highly specialized tissue characterized by an extensive matrix structure formed of fibrous bundles of the protein collagen, and proteoglycans, noncollagenous proteins, lipids and acidic proteins. The processes of bone formation and renewal/repair of bone tissue, which occur continuously throughout life, are performed by specialized cells. Normal embryonic long bone development is preceded by formation of a cartilage model. Bone growth is presumably mediated by "osteoblasts" (bone-forming cells), while remodeling of bone is apparently accomplished by the joint activities of bone-resorbing cells, called "osteoclasts" and osteoblasts. A variety of osteogenic, cartilage-inducing and bone inducing factors have been described. See, e.g. European patent applications 148,155 and 169,016 for discussions thereof.

Brief Description of the Invention

The present invention provides novel proteins in purified form. Specifically, four of the novel proteins are designated BMP-1, BMP-2 Class I (or BMP-2), BMP-3, and BMP-2 Class II (or BMP-4) wherein BMP is bone morphogenic protein. These proteins are characterized by peptide sequences the same as or substantially homologous to amino acid sequences illustrated in Tables II through VIII below. They are capable of inducing bone formation at a predetermined site. These bone inductive factors are further characterized by biochemical and biological characteristics including activity at a concentration of 10 to 1000ng/gram of bone in an *in vivo* rat bone formation assay described below. Proteins of this invention may be encoded by the DNA sequences depicted in the Tables or by sequences capable

substantially the same DNA sequence, as depicted in nucleotide #356 through nucleotide #1543 of Table VII. The human peptide sequence identified in Table VII is 396 amino acids in length. hBMP-2 or related bone inductive proteins may also be characterized by at least a portion of this peptide sequence. hBMP-2 Class I is further characterized by the ability to induce bone formation.

The homologous bovine bone inductive protein of the invention designated bBMP-2 Class I (or bBMP-2), has a DNA sequence identified in Table III below which represents the genomic sequence. This bovine DNA sequence has a prospective 129 amino acid coding sequence followed by approximately 205 nucleotides (a presumptive 3' non-coding sequence). bBMP-2, Class I is further characterized by the ability to induce bone formation. A further bone inductive protein composition of the invention is designated BMP-2 Class II or BMP-4. The human protein hBMP-2 Class II (or hBMP-4) is characterized by at least a portion of the same or substantially the same peptide sequence between amino acid #1 through amino acid #408 of Table VIII, which represents the cDNA of hBMP-2 Class II. This peptide sequence is encoded by at least a portion of the same or substantially the same DNA sequence as depicted in nucleotide #403 through nucleotide #1626 of Table VIII. This factor is further characterized by the ability to induce bone formation.

Still another bone inductive factor of the invention, BMP-3, is represented by the bovine homolog bBMP-3. bBMP-3 is characterized by the DNA sequence and amino acid sequence of Table IV A and B which represents the bovine genomic sequence. It is characterized by at least a portion of a peptide sequence the same or substantially the same as amino acid #1 through amino acid #175 of Table IV A and B. BMP-3 is further characterized by the ability to induce bone formation. The bovine factor may be employed as a tool for obtaining the analogous human BMP-3 protein or other mammalian bone inductive proteins. The proper characterization of this bovine bone

sequences of Tables II through VIII, whether such nucleotide changes result in changes in the peptide sequence or not, are also included in the present invention.

Still a further aspect of the invention is a vector containing a DNA sequence as described above in operative association with an expression control sequence. Such vector may be employed in a novel process for producing a bone growth factor polypeptide in which a cell line transformed with a DNA sequence encoding expression of a bone growth factor polypeptide in operative association with an expression control sequence therefor, is cultured. This claimed process may employ a number of known cells as host cells for expression of the polypeptide. Presently preferred cell lines are mammalian cell lines and bacterial cells.

Other aspects and advantages of the present invention will be apparent upon consideration of the following detailed description and preferred embodiments thereof.

Detailed Description of the Invention

The proteins of the present invention are characterized by amino acid sequences or portions thereof the same as or substantially homologous to the sequences shown in Tables II - VIII below. These proteins are also characterized by the ability to induce bone formation.

The bone growth factors provided herein also include factors encoded by the sequences similar to those of Tables II - VIII, but into which modifications are naturally provided (e.g. allelic variations in the nucleotide sequence which may result in amino acid changes in the polypeptide) or deliberately engineered. For example, synthetic polypeptides may wholly or partially duplicate continuous sequences of the amino acid residues of Tables II - VIII. These sequences, by virtue of sharing primary, secondary, or tertiary structural and conformational characteristics with bone growth factor polypeptides of Tables II - VIII may possess bone growth factor biological properties in common therewith. Thus, they may be

pg. 6 ?

or disulfide linkages, with the sequences of Tables II - VIII and encodes a bone growth factor having one or more bone growth factor biological properties clearly encodes a member of this novel family of growth factors, even if such a DNA sequence would not stringently hybridize to the sequence of Tables II - VIII.

Similarly, DNA sequences which code for bone growth factor polypeptides coded for by the sequences of Tables II - VIII, but which differ in codon sequence due to the degeneracies of the genetic code or allelic variations (naturally-occurring base changes in the species population which may or may not result in an amino acid change) also encode the novel growth factors described herein. Variations in the DNA sequences of Tables II - VIII which are caused by point mutations or by induced modifications to enhance the activity, half-life or production of the polypeptides encoded thereby are also encompassed in the invention.

Another aspect of the present invention provides a novel method for producing the novel osteoinductive factors. The method of the present invention involves culturing a suitable cell or cell line, which has been transformed with a DNA sequence coding on expression for a novel bone growth factor polypeptide of the invention, under the control of known regulatory sequences. Suitable cells or cell lines may be mammalian cells, such as Chinese hamster ovary cells (CHO). The selection of suitable mammalian host cells and methods for transformation, culture, amplification, screening and product production and purification are known in the art. See, e.g., Gething and Sambrook, Nature, 293:620-625 (1981), or alternatively, Kaufman et al, Mol. Cell. Biol., 5(7):1750-1759 (1985) or Howley et al, U.S. Patent 4,419,446. Another suitable mammalian cell line, which is described in the accompanying examples, is the monkey COS-1 cell line. A similarly useful mammalian cell line is the CV-1 cell line.

Bacterial cells are suitable hosts. For example, the

fracture reduction and also in the improved fixation of artificial joints. De novo bone formation induced by an osteogenic agent contributes to the repair of congenital, trauma induced, or oncologic resection induced craniofacial defects, and also is useful in cosmetic plastic surgery. An osteogenic factor of the invention may be valuable in the treatment of periodontal disease, and in other tooth repair processes. Such agents may provide an environment to attract bone-forming cells, stimulate growth of bone-forming cells or induce differentiation of progenitors of bone-forming cells. Of course, the proteins of the invention may have other therapeutic uses.

A further aspect of the invention is a therapeutic method and composition for repairing fractures and other conditions related to bone defects or periodontal diseases. Such a composition comprises a therapeutically effective amount of at least one of the bone inductive factor proteins of the invention. The bone inductive factors according to the present invention may be present in a therapeutic composition in admixture with a pharmaceutically acceptable vehicle or matrix. Further therapeutic methods and compositions of the invention comprise a therapeutic amount of a bone inductive factor of the invention with a therapeutic amount of at least one of the other bone inductive factors of the invention. Additionally, the proteins according to the present invention or a combination of the proteins of the present invention may be co-administered with one or more different osteoinductive factors with which it may interact. Further, the bone inductive proteins may be combined with other agents beneficial to the treatment of the bone defect in question. Such agents include, but are not limited to various growth factors. The preparation of such physiologically acceptable protein compositions, having due regard to pH, isotonicity, stability and the like, is within the skill of the art.

In particular, BMP-1 may be used individually in a

applications.

The choice of material is based on, for example, biocompatibility, biodegradability, mechanical properties, cosmetic appearance and interface properties. Similarly, the application of the osteoinductive factors will define the appropriate formulation. Potential matrices for the osteoinductive factors may be biodegradable and chemically defined, such as, but not limited to calcium sulfate, tricalciumphosphate, hydroxyapatite, polylactic acid, polyanhydrides; biodegradable and biologically well defined, such as bone or dermal collagen, other pure proteins or extracellular matrix components; nonbiodegradable and chemically defined, such as sintered hydroxyapatite, bioglass, aluminates, or other ceramics; or combinations of any of the above mentioned types of material, such as polylactic acid and hydroxyapatite or collagen and tricalciumphosphate. The bioceramics might also be altered in composition, such as in calcium-aluminate-phosphate and processing to alter for example, pore size, particle size, particle shape, and biodegradability.

The dosage regimen will be determined by the attending physician considering various factors which modify the action of such a growth factor, e.g. amount of bone weight desired to be formed, the site of bone damage, the condition of the damaged bone, the patient's age, sex, and diet, the severity of any infection, time of administration and other clinical factors. The dosage may vary with the type of matrix used in the reconstitution and the composition of BMP's. The addition of other known growth factors, such as IGF 1 (insulin like growth factor 1), to the final composition, may also effect the dosage. Generally, the dosage regimen should be in the range of approximately 10 to 10^6 nanograms of protein per gram of bone weight desired. Progress can be monitored by periodic assessment of bone growth and/or repair, e.g. x-rays. Such therapeutic compositions are also presently valuable for veterinary applications due to the lack of species specificity

are collected.

The unbound fractions are concentrated and dialyzed against 50mM NaAc, 50mM NaCl (pH 4.6) in 6M urea. The unbound fractions are applied to a carboxymethyl cellulose column. Protein not bound to the column is removed by extensive washing with starting buffer, and the bone inductive factor containing material desorbed from the column by 50mM NaAc, 0.25mM NaCl, 6M urea (pH 4.6). The protein from this step elution is concentrated 20- to 40- fold, then diluted 5 times with 80mM KPO_4, 6M urea (pH6.0). The pH of the solution is adjusted to 6.0 with 500mM K_2HPO_4. The sample is applied to an hydroxylapatite column (LKB) equilibrated in 80mM KPO_4, 6M urea (pH6.0) and all unbound protein is removed by washing the column with the same buffer. Bone inductive factor activity is eluted with 100mM KPO_4 (pH7.4) and 6M urea. The protein is concentrated approximately 10 times, and solid NaCl added to a final concentration of 0.15M. This material is applied to a heparin - Sepharose column equilibrated in 50mM KPO_4, 150mM NaCl, 6M urea (pH7.4). After extensive washing of the column with starting buffer, a protein with bone inductive factor activity is eluted by 50mM KPO_4, 700mM NaCl, 6M urea (pH7.4). This fraction is concentrated to a minimum volume, and 0.4ml aliquots are applied to Superose 6 and Superose 12 columns connected in series, equilibrated with 4M GuCl, 20mM Tris (pH7.2) and the columns developed at a flow rate of 0.25ml/min. The protein demonstrating bone inductive factor activity has a relative migration corresponding to approximately 30,000 dalton protein.

The above fractions are pooled, dialyzed against 50mM NaAc, 6M urea (pH4.6), and applied to a Pharmacia MonoS HR column. The column is developed with a gradient to 1.0M NaCl, 50mM NaAc, 6M urea (pH4.6). Active fractions are pooled and brought to pH3.0 with 10% trifluoroacetic acid (TFA). The material is applied to a 0.46 x 25cm Vydac C4 column in 0.1% TFA and the column developed with a gradient to 90% acetonitrile, 0.1% TFA

B. Isoelectric Focusing

The isoelectric point of bone inductive factor activity is determined in a denaturing isoelectric focusing system. The Triton X100 urea gel system (Hoeffer Scientific) is modified as follows: 1) 40% of the ampholytes used are Servalyte 3/10; 60% are Servalyte 7-9. 2) The catholyte used is 40mM NaOH. Approximately 20ug of protein from Example I is lyophilized, dissolved in sample buffer and applied to the isoelectrofocusing gel. The gel is run at 20 watts, 10°C for approximately 3 hours. At completion the lane containing bone inductive factor is sliced into 0.5 cm slices. Each piece is mashed in 1.0ml 6M urea, 5mM Tris (pH 7.8) and the samples agitated at room temperature. The samples are acidified, filtered, desalted and assayed as described above. The major portion of activity as determined in the assay described in Example III migrates in a manner consistent with a pI of 8.8 and 9.2.

C. Subunit Characterization

The subunit composition of bone inductive factor is also determined. Pure bone inductive factor is isolated from a preparative 15% SDS gel as described above. A portion of the sample is then reduced with 5mM DTT in sample buffer and re-electrophoresed on a 15% SDS gel. The approximately 30kd protein yields two major bands at approximately 20kd and 18kd, as well as a minor band at 30kd. The broadness of the two bands indicates heterogeneity caused most probably by glycosylation, other post-translational modification, proteolytic degradation or carbamylation.

EXAMPLE III

Biological Activity of Bone Inductive Factor

A rat bone formation assay according to the general procedure of Sampath and Reddi, Proc. Natl. Acad. Sci. U.S.A.,

above, followed by autoradiography. Analysis reveals a correlation of activity with protein bands at 28 - 30kd and a pI 9.0. An extinction coefficient of 1 OD/mg-cm is used as an estimate for protein and approximating the purity of bone inductive factor in a particular fraction. In the in vivo rat bone formation assays on dilutions as described above, the protein is active in vivo at 10 to 200ng protein/gram bone to probably greater than lug protein/gram bone.

EXAMPLE IV

Bovine Bone Inductive Factor Protein Composition

The protein composition of Example IIA of molecular weight 28 - 30kd is reduced as described in Example IIC and digested with trypsin. Eight tryptic fragments are isolated by standard procedures having the following amino acid sequences:

Fragment 1: A A P L G D I A L D E E D L G

Fragment 2: A F Q V Q Q A A D L

Fragment 3: N Y Q D M V V E G

Fragment 4: S T P A Q D V S R

Fragment 5: N Q E A L R

Fragment 6: L S E P D P S H T L E E

Fragment 7: F D A Y Y

Fragment 8: L K P S N ? A T I Q S I V E

Sample A less highly purified preparation of protein from bovine bone is prepared according to a purification scheme similary to that described in Example I. The purification basically varies from that previously described by omission of the DE-52 column, the CM cellulose column and the mono S column, as well as a reversal in the order of the hydroxylapatite and heparin sepharose columns. Briefly, the concentrted crude 4 M extract is brought to 85% final concentration of ethanol at 4 degrees. The mixture is then centrifuged, and the precipitate redissolved in 50 mM Tris, 0.15 M NaCl, 6.0 M urea. This material is then fractionated on Heparin Sepharose as described. The Heparin bound material

(a) A [A/G] [A/G] TC [T/C] TC [T/C] TC [A/G] TC [T/C] AA

(b) A [A/G] [A/G] TC [T/C] TC [T/C] TC [A/G] TCNAG

Bracketed nucleotides are alternatives. "N" means either A, T, C or G.

In both cases the regions of the amino acid sequence used for probe design are chosen by avoiding highly degenerate codons where possible. The oligonucleotides are synthesized on an automated DNA synthesizer; the probes are then radioactively labeled with polynucleotide kinase and ^{32}P-ATP.

These two sets of probes are used to screen a bovine genomic recombinant library. The library is constructed as follows: Bovine liver DNA is partially digested with the restriction endonuclease enzyme Sau 3A and sedimented through a sucrose gradient. Size fractionated DNA in the range of 15-30kb is then ligated to the bacteriophage Bam HI vector EMBL3 [Frischauf et al, J. Mol. Biol., 170:827-842 (1983)]. The library is plated at 8000 recombinants per plate. Duplicate nitrocellulose replicas of the plaques are made and amplified according to a modification of the procedure of Woo et al, Proc. Natl. Acad. Sci. USA, 75:3688-91 (1978).

The 32 mer probe is kinased with ^{32}P-gamma-ATP and hybridized to one set of filters in 5X SSC, 0.1% SDS, 5X Denhardts, 100ug/ml salmon sperm DNA at 45 degrees C and washed with 5X SSC, 0.1% SDS at 45 degrees C. The 17 mer probes are kinased and hybridized to the other set of filters in 3M tetramethylammonium chloride (TMAC), 0.1M sodium phosphate pH6.5, 1mM EDTA, 5X Denhardts, 0.6% SDS, 100ug/ml salmon sperm DNA at 48 degrees C, and washed in 3M TMAC, 50mM Tris pH8.0 at 50 degrees C. These conditions minimize the detection of mismatches to the 17 mer probe pool [see, Wood et al, Proc. Natl. Acad. Sci, U.S.A., 82:1585-1588 (1985)]. 400,000 recombinants are screened by this procedure and one duplicate positive is plaque purified. DNA is isolated from a plate lysate of this recombinant bacteriophage designated lambda bP-50. bP-50 was deposited December 16, 1986 with the American

where is pg. 20 ?

TABLE II

280 290 (1) 308 323

CCTTGCCTCT TCTCTCTCCA GCT GCC TTC CTT GGG GAC ATC GCC CTG GAC GAG GAG

Ala Phe Leu Gly Asp Ile Ala Leu Asp Glu Glu

A F L G D I A L D E E

338 353 368

GAC TTG AGG GCC TTC CAA GTG CAG CAG GCT GCG GAC CTC AGA CAG CGT GCA ACC

Asp Leu Arg Ala Phe Gln Val Gln Gln Ala Ala Asp Leu Arg Gln Arg Ala Thr

D L R A P Q V Q Q A A D L R Q R A T

383 398 (37) 414 424

CGC AGG TCT TCC ATC AAA GCT GCA GGTACACTGG GTACAGGCCA

Arg Arg Ser Ser Ile Lys Ala Ala

R R S S I K A A

residue (K) as expected considering the specificity of trypsin. The arginine residue encoded by the CGT triplet is presumed to be the carboxy-terminus of the protein based on the presence of a stop codon (TAG) adjacent to it.

GAC TTG AGG GCC TTC CAA GTG CAG CAG GCT GCT TAG CTG AG CGT GCA AC

Cas bBMP-3 [illegible]

The arg Probes consisting of pools of oligonucleotides are designed on the basis of the amino acid sequences of the tryptic Fragments 9 (Probe #3), 10 (Probe #2), and 11 (Probe #1), and synthesized on an automated DNA synthesizer.

Probe #1: A C N G T C A T [A/G] T T N G G [A/G] T A

Probe #2: C A [A/G] T A [A/G] T A N G C [A/G] T C [A/G] A A

Probe #3: T G [A/G/T] A T N G T N G C [A/G] T G [A/G] T T

A recombinant bovine genomic library constructed in EMBL3 is screened by the TMAC hybridization procedure detailed above in part A. 400,000 recombinants are screened in duplicate with Probe #1 which has been labeled with ^{32}P. All recombinants which hybridized to this probe are replated for secondaries. Triplicate nitrocellulose replicas are made of the secondary plates, and amplified as described. The three sets of filters are hybridized to Probes #1, #2 and #3, again under TMAC conditions. One clone, lambda bP-819, hybridizes to all three probes and is plaque purified and DNA is isolated from a plate lysate. Bacteriophage lambda bP-819 was deposited with the ATCC on June 16, 1987 under accession number 40344. This bP-819 clone encodes the bovine bone growth factor designated bBMP-3.

The region of bP-819 which hybridizes to Probe #2 is localized and sequenced. The partial DNA and derived amino acid sequences of this region are shown in Table IVA. The amino acid sequences corresponding to tryptic Fragments 10 and 12 are underlined. The first underlined sequence corresponds to Fragment 12 while the second corresponds to Fragment 10. This region of bP-819, therefore, which hybridizes to Probe #2 encodes at least 111 amino acids. This amino acid sequence is encoded by the DNA sequence from nucleotide #414 through #746.

EXAMPLE V

Human Bone Inductive Factors

A. hBMP-1

Because the bovine and human bone growth factor genes are presumed to be significantly homologous, the bovine bBMP-1 DNA sequence of Table II (or portions thereof) is used as a probe to screen a human genomic library. The 800bp EcoRI fragment of the bovine genomic clone is labeled with ^{32}P by nick-translation. A human genomic library (Toole et al., supra) is plated on 20 plates at 40,000 recombinants per plate. Duplicate nitrocellulose filter replicas are made of each plate and hybridized to the nick-translated probe in 5 X SSC, 5 X Denhardt's, 100ug/ml denatured salmon sperm DNA, 0.1% SDS (the standard hybridization solution) at 50 degrees centigrade for approximately 14 hours. The filters are then washed in 1 X SSC, 0.1% SDS at 50 degrees centigrade and subjected to autoradiography. Five duplicate positives are isolated and plaque purified. DNA is obtained from a plate lysate of one of these recombinant bacteriophage, designated LP-H1. LP-H1 was deposited with the ATCC on March 6, 1987 under accession number 40311. This clone encodes at least a portion of the human genomic bone growth factor called hBMP-1. The hybridizing region of LP-H1 is localized to a 2.5kb XbaI/HindIII restriction fragment.

The partial DNA sequence and derived amino acid sequence of lambda LP-H1 are shown below in Table V. The peptide sequence from this clone is 37 amino acids in length and is encoded by the DNA sequence from nucleotide #3440 through nucleotide #3550. The coding sequence of Table V is flanked by approximately 28 nucleotides (a presumptive 5' noncoding sequence) as well as approximately 19 nucleotides (a presumptive 3' noncoding sequence. A comparison of the bBMP-1 sequence of Table II with the hBMP-1 genomic sequence of Table V indicates the significant homology between the two.

Because the size of coding regions and the positions

Human Bone Inductive Factors TABLE V

```
3419       3429       3439       (1)             3454
CAGCCCTGGC TTCTTCTTTT CTCTTTAGCT GCC TTT CTT GGG GAC ATT GCC CTG GAC
                                 Ala Phe Leu Gly Asp Ile Ala Leu Asp
                                 A   F   L   G   D   I   A   L   D
3469             3484            3499            3514
GAA GAG GAC CTG AGG GCC TTC CAG GTA CAG CAG GCT GTG GAT CTC AGA CGG CAC
Glu Glu Asp Leu Arg Ala Phe Gln Val Gln Gln Ala Val Asp Leu Arg Arg His
E   E   D   L   R   A   F   Q   V   Q   Q   A   V   D   L   R   R   H
    3529            3544   (37)     3560       3570
ACA GCT CGT AAG TCC TCC ATC AAA GCT GCA GGTAAGCCGG GTGCCAATGG
Thr Ala Arg Lys Ser Ser Ile Lys Ala Ala
T   A   R   L   S   S   I   L   A   A
```

[illegible] each plate and hybridized to the nick-translated [illegible] SSC, 5 X Denhardt's, [illegible] denatured salmon sperm [illegible] 0.1% SDS (the standard hybridization solution) [illegible] degrees centigrade for approximately [illegible] hours. [illegible] washed in 1 X SSC, 0.1% SDS at 50 degrees centigrade and subjected to autoradiography. Five duplicate [illegible] isolated and plaque purified. DNA is obtained [illegible] lysate of one of these recombinant bacteriophage [illegible] LP-H1. LP-H1 was deposited with the ATCC on March [illegible] under accession number 40311. This clone encodes at least a portion of the human genomic bone growth factor called [illegible] hybridizing region of LP-H1 is [illegible] restriction fragment.

[illegible] sequence of lambda LP-H1 are shown [illegible] peptide sequence from this clone [illegible] is encoded [illegible] nucleotide [illegible] the coding sequence [illegible] approximately [illegible] coding sequence [illegible]

where is pg. 30?

The entire nucleotide sequence and derived amino acid sequence of the insert of lambda U2OS-1 is given in Table VI. This cDNA clone encodes a Met followed by a hydrophobic leader sequence characteristic of a secreted protein, and contains a stop codon at nucleotide positons 2226 - 2228. This clone contains an open reading frame of 2190bp, encoding a protein of 730 amino acids with a molecular weight og 83kd based on this amino acid sequence. The clone contains sequence identical to the coding region given in Table V. This protein is contemplated to represent a primary translation product which is cleaved upon secretion to produce the hBMP-1 protein. This clone is therefore a cDNA for hBMP-1 corresponding to human gene fragment contained in the genomic hBMP-1 sequence lambda LP-H1. It is noted that amino acids #550 to #590 of BMP-1 are homologous to epidermal growth factor and the "growth factor" domains of Protein C, Factor X and Factor IX.

pg. 32 ?

Pg. 34?

The nucleotide sequence and derived amino acid

```
                665                 680                 695
AAG TTC GGC ATT GTG GTC CAC GAG CTG GGC CAC GTC GTC GGC TTC TGG CAC GAA
Lys Phe Gly Ile Val Val His Glu Leu Gly His Val Val Gly Phe Trp His Glu
K  F  G  I  V  V  H  E  L  G  H  V  V  G  F  W  H  E
710                 725                 740                 755
CAC ACT CGG CCA GAC CGG GAC CGC CAC GTT TCC ATC GTT CGT GAG AAC ATC CAG
His Thr Arg Pro Asp Arg Asp Arg His Val Ser Ile Val Arg Glu Asn Ile Gln
H  T  R  P  D  R  D  R  H  V  S  I  V  R  E  N  I  Q
            770                 785                 800                 815
CCA GGG CAG GAG TAT AAC TTC CTG AAG ATG GAG CCT CAG GAG GTG GAG TCC CTG
Pro Gly Gln Glu Tyr Asn Phe Leu Lys MET Glu Pro Gln Glu Val Glu Ser Leu
P  G  Q  E  Y  N  F  L  K  M  E  P  Q  E  V  E  S  L
                        830                 845                 860
GGG GAG ACC TAT GAC TTC GAC AGC ATC ATG CAT TAC GCT CGG AAC ACA TTC TCC
Gly Glu Thr Tyr Asp Phe Asp Ser Ile MET His Tyr Ala Arg Asn Thr Phe Ser
G  E  T  Y  D  F  D  S  I  M  H  Y  A  R  N  T  F  S
   875                 890                 905                 920
AGG GGC ATC TTC CTG GAT ACC ATT GTC CCC AAG TAT GAG GTG AAC GGG GTG AAA
Arg Gly Ile Phe Leu Asp Thr Ile Val Pro Lys Tyr Glu Val Asn Gly Val Lys
R  G  I  F  L  D  T  I  V  P  K  Y  E  V  N  G  V  K
                  935                 950                 965
CCT CCC ATT GGC CAA AGG ACA CGG CTC AGC AAG GGG GAC ATT GCC CAA GCG CGC
Pro Pro Ile Gly Gln Arg Thr Arg Leu Ser Lys Gly Asp Ile Ala Gln Ala Arg
P  P  I  G  Q  R  T  R  L  S  K  G  D  I  A  Q  A  R
980                 995                 1010                1025
AAG CTT TAC AAG TGC CCA GCC TGT GGA GAG ACC CTG CAA GAC AGC ACA GGC AAC
Lys Leu Tyr Lys Cys Pro Ala Cys Gly Glu Thr Leu Gln Asp Ser Thr Gly Asn
K  L  Y  K  C  P  A  C  G  E  T  L  Q  D  S  T  G  N
      1040                1055                1070                1085
TTC TCC TCC CCT GAA TAC CCC AAT GGC TAC TCT GCT CAC ATG CAC TGC GTG TGG
Phe Ser Ser Pro Glu Tyr Pro Asn Gly Tyr Ser Ala His MET His Cys Val Trp
F  S  S  P  E  Y  P  N  G  Y  S  A  H  M  H  C  V  W
                  1100                1115                1130
CGC ATC TCT GTC ACA CCC GGG GAG AAG ATC ATC CTG AAC TTC ACG TCC CTG GAC
Arg Ile Ser Val Thr Pro Gly Glu Lys Ile Ile Leu Asn Phe Thr Ser Leu Asp
R  I  S  V  T  P  G  E  K  I  I  L  N  F  T  S  L  D
   1145                1160                1175                1190
CTG TAC CGC AGC CGC CTG TGC TGG TAC GAC TAT GTG GAG GTC CGA GAT GGC TTC
Leu Tyr Arg Ser Arg Leu Cys Trp Tyr Asp Tyr Val Glu Val Arg Asp Gly Phe
L  Y  R  S  R  L  C  W  Y  D  Y  V  E  V  R  D  G  F
            1205                1220                1235
TGG AGG AAG GCG CCC CTC CGA GGC CGC TTC TGC GGG TCC AAA CTC CCT GAG CCT
Trp Arg Lys Ala Pro Leu Arg Gly Arg Phe Cys Gly Ser Lys Leu Pro Glu Pro
W  R  K  A  P  L  R  G  R  F  C  G  S  K  L  P  E  P
1250                1265                1280                1295
ATC GTC TCC ACT GAC AGC CGC CTC TGG GTT GAA TTC CGC AGC AGC AGC AAT TGG
Ile Val Ser Thr Asp Ser Arg Leu Trp Val Glu Phe Arg Ser Ser Ser Asn Trp
I  V  S  T  D  S  R  L  W  V  E  F  R  S  S  S  N  W
      1310                1325                1340                1355
GTT GGA AAG GGC TTC TTT GCA GTC TAC GAA GCC ATC TGC GGG GGT GAT GTG AAA
Val Gly Lys Gly Phe Phe Ala Val Tyr Glu Ala Ile Cys Gly Gly Asp Val Lys
W  G  K  G  F  F  A  V  Y  E  A  I  C  G  G  D  V  K
```

pg 36

2060 2075 2090 2105
ACC TCC CAG TAC AAC AAC ATG CGC GTG GAG TTC AAG TCC GAC AAC ACC GTG TCC
Thr Ser Gln Tyr Asn Asn MET Arg Val Glu Phe Lys Ser Asp Asn Thr Val Ser
T S Q Y N N M R V E F K S D N T V S

2120 2135 2150 2165
AAA AAG GGC TTC AAG GCC CAC TTC TTC TCA GAA AAG AGG CCA GCT CTG CAG CCC
Lys Lys Gly Phe Lys Ala His Phe Phe Ser Glu Lys Arg Pro Ala Leu Gln Pro
K K G F K A H F F S E K R P A L Q P

2180 2195 2210
CCT CGG GGA CGC CCC CAC CAG CTC AAA TTC CGA GTG CAG AAA AGA AAC CGG ACC
Pro Arg Gly Arg Pro His Gln Leu Lys Phe Arg Val Gln Lys Arg Asn Arg Thr
P R(730) G R P H Q L K F R V Q K R N R T

2225 2235 2245 2255 2265 2275 2285
CCC CAG TGAGGCCTGC CAGGCCTCCC GGACCCCTTG TTACTCAGGA ACCTCACCTT GGACGGAATG
Pro Gln
P Q

2295 2305 2315 2325 2335 2345 2355
GGATGGGGGC TTCGGTGCCC ACCAACCCCC CACCTCCACT CTGCCATTCC GGCCCACCTC CCTCTGGCCG

2365 2375 2385 2395 2405 2415 2425
GACAGAACTG GTGCTCTCTT CTCCCCACTG TGCCCGTCCG CGGACCGGGG ACCCTTCCCC GTGCCCTACC

2435 2445 2455 2465 2475 2485 2495
CCCTCCCATT TTGATGGTGT CTGTGACATT TCCTGTTGTG AAGTAAAAGA GGGACCCCTG CGTCCTGCCT

CTAGA

with the sequence given in Table III at the 3' end of their coding regions, but less so in the more 5' regions. Thus they encode a human protein of similar, though not identical, structure to that above.

Full length hBMP-2 Class I cDNA clones are obtained in the following manner. The 1.5 kb insert of one of the Class II subclones (II-10-1) is isolated and radioactively labeled by nick-translation. One set of the nitrocellulose replicas of the U-2 OS cDNA library screened above (50 filters, corresponding to 1,000,000 recombinant bacteriophage) are rehybridized with this probe under stringent conditions (hybridization at 65° in standard hybridization buffer; washing at 65° in 0.2 X SSC, 0.1% SDS). All recombinants which hybridize to the bovine genomic probe which do not hybridize to the Class II probe are picked and plaque purified (10 recombinants). Plate stocks are made and small scale bacteriophage DNA preparations made. After subcloning into M13, sequence analysis indicates that 4 of these represent clones which overlap the original Class I clone. One of these, lambda U20S-39, contains an approximately 1.5 kb insert and was deposited with the ATCC on June 16, 1987 under accession number 40345. The partial DNA sequence (compiled from lambda U20S-39 and several other hBMP-2 Class I cDNA recombinants) and derived amino acid sequence are shown below in Table VII. Lambda U20S-39 is expected to contain all of the nucleotide sequence necessary to encode the entire human counterpart of the protein BMP-2 Class II encoded by the bovine gene segment whose partial sequence is presented in Table III. This human cDNA hBMP-2 Class II contains an open reading frame of 1188 bp, encoding a protein of 396 amino acids. This protein of 396 amino acids has a molecular weight of 45kd based on this amino acid sequence. It is contemplated that this sequence represents the primary translation product. The protein is preceded by a 5' untranslated region of 342 bp with stop codons in all frames.

pg 40?

TABLE VII

```
          10         20         30         40         50         60         70
GTCGACTCTA GAGTGTGTGT CAGCACTTGG CTGGGGACTT CTTGAACTTG CAGGGAGAAT AACTTGCGCA

          80         90        100        110        120        130        140
CCCCACTTTG CGCCGGTGCC TTTGCCCCAG CGGAGCCTGC TTCGCCATCT CCGAGCCCCA CCGCCCCTCC

         150        160        170        180        190        200        210
ACTCCTCGGC CTTGCCCGAC ACTGAGACGC TGTTCCCAGC GTGAAAAGAG AGACTGCGCG GCCGGCACCC

         220        230        240        250        260        270        280
GGGAGAAGGA GGAGGCAAAG AAAAGGAACG GACATTCGGT CCTTGCGCCA GGTCCTTTGA CCAGAGTTTT

         290        300        310        320        330        340        350
CCTCCATGTGGA CGCTCTTTCA ATGGACGTGT CCCCGCGTGC TTCTTAGACG GACTGCGGTC TCCTAAAGGT

                   370                 385                 400
CGACC ATG GTG GCC GGG ACC CGC TGT CTT CTA GCG TTG CTG CTT CCC CAG GTC
      MET Val Ala Gly Thr Arg Cys Leu Leu Ala Leu Leu Leu Pro Gln Val
      M  V  A  G  T  R  C  L  L  A  L  L  L  P  Q  V
             415                 430                 445
CTC CTG GGC GGC GCG GCT GGC CTC GTT CCG GAG CTG GGC CGC AGG AAG TTC GCG
Leu Leu Gly Gly Ala Ala Gly Leu Val Pro Glu Leu Gly Arg Arg Lys Phe Ala
 L  L  G  G  A  A  G  L  V  P  E  L  G  R  R  K  F  A
460                 475                 490                 505
GCG GCG TCG TCG GGC CGC CCC TCA TCC CAG CCC TCT GAC GAG GTC CTG AGC GAG
Ala Ala Ser Ser Gly Arg Pro Ser Ser Gln Pro Ser Asp Glu Val Leu Ser Glu
 A  A  S  S  G  R  P  S  S  Q  P  S  D  E  V  L  S  E
        520                 535                 550                 565
TTC GAG TTG CGG CTG CTC AGC ATG TTC GGC CTG AAA CAG AGA CCC ACC CCC AGC
Phe Glu Leu Arg Leu Leu Ser MET Phe Gly Leu Lys Gln Arg Pro Thr Pro Ser
 F  E  L  R  L  L  S  M  F  G  L  K  Q  R  P  T  P  S
                580                 595                 610
AGG GAC GCC GTG GTG CCC CCC TAC ATG CTA GAC CTG TAT CGC AGG CAC TCG GGT
Arg Asp Ala Val Val Pro Pro Tyr MET Leu Asp Leu Tyr Arg Arg His Ser Gly
 R  D  A  V  V  P  P  Y  M  L  D  L  Y  R  R  H  S  G
625                 640                 655                 670
CAG CCG GGC TCA CCC GCC CCA GAC CAC CGG TTG GAG AGG GCA GCC AGC CGA GCC
Gln Pro Gly Ser Pro Ala Pro Asp His Arg Leu Glu Arg Ala Ala Ser Arg Ala
 Q  P  G  S  P  A  P  D  H  R  L  E  R  A  A  S  R  A
685                 700                 715
AAC ACT GTG CGC AGC TTC CAC CAT GAA GAA TCT TTG GAA GAA CTA CCA GAA ACG
Asn Thr Val Arg Ser Phe His His Glu Glu Ser Leu Glu Glu Leu Pro Glu Thr
 N  T  V  R  S  F  H  H  E  E  S  L  E  E  L  P  E  T
```

```
   1435                1450                1465                1480
GCA TGC TGT GTC CCG ACA GAA CTC AGT GCT ATC TCG ATG CTG TAC CTT GAC GAG
Ala Cys Cys Val Pro Thr Glu Leu Ser Ala Ile Ser MET Leu Tyr Leu Asp Glu
 A   C   C   V   P   T   E   L   S   A   I   S   M   L   Y   L   D   E
            1495                1510                1525
AAT GAA AAG GTT GTA TTA AAG AAC TAT CAG GAC ATG GTT GTG GAG GGT TGT GGG
Asn Glu Lys Val Val Leu Lys Asn Tyr Gln Asp MET Val Val Glu Gly Cys Gly
 N   E   K   V   V   L   K   N   Y   Q   D   M   V   V   E   G   C   G
1540(396)    1553        1563        1573        1583        1593        1603
TGT CGC TAGTACAGCA AAATTAAATA CATAAATATA TATATATATA TATATTTTAG AAAAAAGAAA
Cys Arg
 C  R

AAAA
```

pg 44

primary translation product.

```
                747                 762                 777
AGC CGG GCC AAC ACC GTG AGG AGC TTC CAC CAC GAA GAA CAT CTG GAG AAC ATC
Ser Arg Ala Asn Thr Val Arg Ser Phe His His Glu Glu His Leu Glu Asn Ile
 S  R  A  N  T  V  R  S  F  H  H  E  E  H  L  E  N  I
792                 807                 822                 837
CCA GGG ACC AGT GAA AAC TCT GCT TTT CGT TTC CTC TTT AAC CTC AGC AGC ATC
Pro Gly Thr Ser Glu Asn Ser Ala Phe Arg Phe Leu Phe Asn Leu Ser Ser Ile
 P  G  T  S  E  N  S  A  F  R  F  L  F  N  L  S  S  I
        852                 867                 882                 897
CCT GAG AAC GAG GTG ATC TCC TCT GCA GAG CTT CGG CTC TTC CGG GAG CAG GTG
Pro Glu Asn Glu Val Ile Ser Ser Ala Glu Leu Arg Leu Phe Arg Glu Gln Val
 P  E  N  E  V  I  S  S  A  E  L  R  L  F  R  E  E  V
                 912                 927                 942
GAC CAG GGC CCT GAT TGG GAA AGG GGC TTC CAC CGT ATA AAC ATT TAT GAG GTT
Asp Gln Gly Pro Asp Trp Glu Arg Gly Phe His Arg Ile Asn Ile Tyr Glu Val
 D  Q  G  P  D  W  E  R  G  F  H  R  I  N  I  Y  E  V
    957                 972                 987                1002
ATG AAG CCC CCA GCA GAA GTG GTG CCT GGG CAC CTC ATC ACA CGA CTA CTG GAC
MET Lys Pro Pro Ala Glu Val Val Pro Gly His Leu Ile Thr Arg Leu Leu Asp
 M  K  P  P  A  E  V  V  P  G  H  L  I  T  R  L  L  D
           1017                1032                1047
ACG AGA CTG GTC CAC CAC AAT GTG ACA CGG TGG GAA ACT TTT GAT GTG AGC CCT
Thr Arg Leu Val His His Asn Val Thr Arg Trp Glu Thr Phe Asp Val Ser Pro
 T  R  L  V  H  H  N  V  T  R  W  E  T  F  D  V  S  P
1062                1077                1092                1107
GCG GTC CTT CGC TGG ACC CGG GAG AAG CAG CCA AAC TAT GGG CTA GCC ATT GAG
Ala Val Leu Arg Trp Thr Arg Glu Lys Gln Pro Asn Tyr Gly Leu Ala Ile Glu
 A  V  L  R  W  T  R  E  K  Q  P  N  Y  G  L  A  I  E
       1122                1137                1152                1167
GTG ACT CAC CTC CAT CAG ACT CGG ACC CAC CAG GGC CAG CAT GTC AGG ATT AGC
Val Thr His Leu His Gln Thr Arg Thr His Gln Gly Gln His Val Arg Ile Ser
 V  T  H  L  H  Q  T  R  T  H  Q  G  Q  H  V  R  I  S
                1182                1197                1212
CGA TCG TTA CCT CAA GGG AGT GGG AAT TGG GCC CAG CTC CGG CCC CTC CTG GTC
Arg Ser Leu Pro Gln Gly Ser Gly Asn Trp Ala Gln Leu Arg Pro Leu Leu Val
 R  S  L  P  Q  G  S  G  N  W  A  Q  L  R  P  L  L  V
   1227                1242                1257                1272
ACC TTT GGC CAT GAT GGC CGG GGC CAT GCC TTG ACC CGA CGC CGG AGG GCC AAG
Thr Phe Gly His Asp Gly Arg Gly His Ala Leu Thr Arg Arg Arg Arg Ala Lys
 T  F  G  H  D  G  R  G  H  A  L  T  R  R  R  R  A  K
          1287                1302                1317
CGT AGC CCT AAG CAT CAC TCA CAG CGG GCC AGG AAG AAG AAT AAG AAC TGC CGG
Arg Ser Pro Lys His His Ser Gln Arg Ala Arg Lys Lys Asn Lys Asn Cys Arg
 R  S  P  K  H  H  S  Q  R  A  R  K  K  N  K  N  C  R
1332                1347                1362                1377
CGC CAC TCG CTC TAT GTG GAC TTC AGC GAT GTG GGC TGG AAT GAC TGG ATT GTG
Arg His Ser Leu Tyr Val Asp Phe Ser Asp Val Gly Trp Asn Asp Trp Ile Val
 R  H  S  L  Y  V  D  F  S  D  V  G  W  N  D  W  I  V
       1392                1407                1422                1437
GCC CCA CCA GGC TAC CAG GCC TTC TAC TGC CAT GGG GAC TGC CCC TTT CCA CTG
Ala Pro Pro Gly Tyr Gln Ala Phe Tyr Cys His Gly Asp Cys Pro Phe Pro Leu
 A  P  P  G  Y  Q  A  F  Y  C  H  G  D  C  P  F  P  L
```

The sequences of BMP-2 Class I and II, as well as BMP-3 as shown in Tables III, IV, VII and VIII have significant homology to the beta (B) and beta (A) subunits of the inhibins. The inhibins are a family of hormones which are presently being investigated for use in contraception. See, A. J. Mason et al, Nature, 318:659-663 (1985). To a lesser extent they are also homologous to Mullerian inhibiting substance (MIS), a testicular glycoprotein that causes regression of the Mullerian duct during development of the male embryo and transforming growth factor-beta (TGF-b) which can inhibit or stimulate growth of cells or cause them to differentiate. Furthermore, the sequence of Table VII encoding hBMP-2 Class II has significant homology to the Drosophila decapentaplegic (DPP-C) locus transcript. See, J. Massague, Cell, 49:437-438 (1987); R. W. Padgett et al, Nature, 325:81-84 (1987); R. L. Cate et al, Cell 45: 685-698 (1986). It is considered possible therefore that BMP-2 Class II is the human homolog of the protein made from this transcript from this developmental mutant locus.

C. BMP-3

Because bovine and human bone growth factor genes are presumed to be significantly homologous, oligonucleotide probes which have been shown to hybridize to the bovine DNA sequence of Table IV.A and IV.B are used to screen a human genomic library. A human genomic library (Toole et al., supra) is screened using these probes, and presumptive positives are isolated and DNA sequence obtained as described above. Evidence that this recombinant encodes a portion of the human bone inductive factor molecule relies on the bovine/human protein and gene structure homologies.

Once a recombinant bacteriophage containing DNA encoding a portion of the human BMP-3 molecule is obtained the human coding sequence is used as a probe as described in Example V (A) to identify a human cell line or tissue which synthesizes BMP-3. mRNA is selected by oligo (dT) cellulose

VIII or other modified sequences and known vectors, such as pCD [Okayama et al., Mol. Cell Biol., 2:161-170 (1982)] and pJL3, pJL4 [Gough et al., EMBO J., 4:645-653 (1985)]. The transformation of these vectors into appropriate host cells can result in expression of osteoinductive factors. One skilled in the art could manipulate the sequences of Tables II-VIII by eliminating or replacing the mammalian regulatory sequences flanking the coding sequence with bacterial sequences to create bacterial vectors for intracellular or extracellular expression by bacterial cells. For example, the coding sequences could be further manipulated (e.g. ligated to other known linkers or modified by deleting non-coding sequences there-from or altering nucleotides therein by other known techniques). The modified bone inductive factor coding sequence could then be inserted into a known bacterial vector using procedures such as described in T. Taniguchi et al., Proc. Natl Acad. Sci. USA, 77:5230-5233 (1980). This exemplary bacterial vector could then be transformed into bacterial host cells and bone inductive factor expressed thereby. For a strategy for producing extracellular expression of bone inductive factor in bacterial cells, see, e.g. European patent application EPA 177,343.

Similar manipulations can be performed for the construction of an insect vector [See, e.g. procedures described in published European patent application 155,476] for expression in insects cells. A yeast vector could also be constructed employing yeast regulatory sequences for intracellular or extracellular expression of the factors of the present invention by yeast cells. [See, e.g., procedures described in published PCT application WO86/00639 and European patent application EPA 123,289].

A method for producing high levels of an osteoinductive factor of the invention from mammalian cells involves the construction of cells containing multiple copies of the heterologous bone inductive factor gene. The heterologous gene

Danna, PNAS 78:7575-7578 (1981); Luthman and Magnusson, Nucl. Acids Res. 11: 1295-1308 (1983)]. Serum-free 24 hr. conditioned medium is collected from the cells starting 40-70 hr. post-transfection.

The mammalian expression vector pMT2 Cla-Xho (pMT2 CX) is a derivative of p91023 (b) (Wong et al., Science 228:810-815, 1985) differing from the latter in that it contains the ampicillin resistance gene in place of the tetracycline resistance gene and further contains a XhoI site for insertion of cDNA clones. The functional elements of pMT2 Cla-Xho have been described (Kaufman, R.J., 1985, Proc. Natl. Acad. Sci. USA 82:689-693) and include the adenovirus VA genes, the SV40 origin of replication including the 72 bp enhancer, the adenovirus major late promoter including a 5' splice site and the majority of the adenovirus tripartite leader sequence present on adenovirus late mRNAs, a 3' splice acceptor site, a DHFR insert, the SV40 early polyadenylation site (SV40), and pBR322 sequences needed for propagation in E. coli.

Plasmid pMT2 Cla-Xho is obtained by EcoRI digestion of pMT2-VWF, which has been deposited with the American Type Culture Collection (ATCC), Rockville, MD (USA) under accession number ATCC 67122. EcoRI digestion excises the cDNA insert present in pMT2-VWF, yielding pMT2 in linear form which can be ligated and used to transform E. coli HB 101 or DH-5 to ampicillin resistance. Plasmid pMT2 DNA can be prepared by conventional methods. pMT2CX is then constructed by digesting pMT2 with Eco RV and XbaI, treating the digested DNA with Klenow fragment of DNA polymerase I, and ligating Cla linkers (NEBiolabs, CATCGATG). This removes bases 2266 to 2421 starting from the Hind III site near the SV40 origin of replication and enhancer sequences of pMT2. Plasmid DNA is then digested with EcoRI, blunted as above, and ligated to an EcoRI adapter,

5' PO_4-AATTCCTCGAGAGCT 3'

Addition of human BMP-1 to the matrix material resulted in formation of cartilage-like nodules at 7 days post implantation. The chondroblast-type cells were recognizable by shape and expression of metachromatic matrix. The amount of activity observed for human BMP-1 was dependent upon the amount of human BMP-1 protein added to the matrix. Table IX illustrates the dose-response relationship of human BMP-1 protein to the amount of bone induction observed.

Table IX

IMPLANT NUMBER	AMOUNT USED (equivalent of ml transfection media)		HISTOLOGICAL SCORE
876-134-1	10	BMP-1	C+2
876-134-2	3	BMP-1	C+1
876-134-3	1	BMP-1	C +/-
876-134-4	10	MOCK	C -
876-134-5	3	MOCK	C -
876-134-6	1	MOCK	C -

Cartilage (c) activity was scored on a scale from 0(-) to 5.

Similar levels of activity are seen in the Heparin Sepharose fractionated COS cell extracts. Partial purification is accomplished in a similar manner as described above except that 6 M urea is included in all the buffers. Further, in a rat bone formation assay as described above, BMP-2 has similarly demonstrated chondrogenic activity.

The procedures described above may be employed to isolate other bone inductive factors of interest by utilizing the bovine bone inductive factors and/or human bone inductive factors as a probe source. Such other bone inductive factors may find similar utility in, inter alia, fracture repair.

The foregoing descriptions detail presently preferred embodiments of the present invention. Numerous modifications

International Application No: PCT/ /

MICROORGANISMS

Optional Sheet in connection with the microorganism referred to on page ____, line ____ of the description [1]

A. IDENTIFICATION OF DEPOSIT [2]

Further deposits are identified on an additional sheet ☐ [3]

Name of depositary institution [4]

amount of human BMP-1 protein added to the medium. Table 2.

American Type Culture Collection

Address of depositary institution (including postal code and country) [5]

12301 Parklawn Drive
Rockville, Maryland 20852 USA

Name of Deposit	ATCC No.	Referred to on page/line	Date of Deposit
LP-H1	40311	29/20	March 4, 1987
bP50	40295	20/3	December 15, 1986
bP-21	40310	22/18	March 4, 1987
U2OS-3	40342	44/22	June 16, 1987
Lambda U2-OS-1	40343	32/33	June 16, 1987
Lambda BP819	40344	25/23	June 16, 1987
U2OS-39	40345.	39/21	June 16, 1987

C. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE [3] (if the indications are not for all designated States)

Cancellation of additional designated States

D. SEPARATE FURNISHING OF INDICATIONS [6] (leave blank if not applicable)

The indications listed below will be submitted to the International Bureau later [6] (Specify the general nature of the indications e.g., "Accession Number of Deposit")

E. ☑ This sheet was received with the international application when filed (to be checked by the receiving Office)

Thadi Singh
(Authorized Officer)

☐ The date of receipt (from the applicant) by the International Bureau [7]

(Authorized Officer)

Form PCT RO 134 (January 1981)

with a DNA sequence encoding BMP-2 Class I, said DNA sequence being in relative association with an expression control sequence therefor, and isolating BMP-2 Class I from said culture medium.

12. A process for according to Claim 11 wherein said DNA sequence comprises substantially the nucleotide sequence of Table VII.

13. A process for producing BMP-2 Class II comprising culturing in a suitable culture medium a cell line transformed with a DNA sequence encoding BMP-2 Class II, said DNA sequence being in relative association with an expression control sequence therefor, and isolating BMP-2 Class II from said culture medium.

14. A process according to Claim 13 wherein said DNA sequence comprises substantially the nucleotide sequence of Table VIII.

15. A process for producing BMP-3 comprising culturing in a suitable culture medium a cell line transformed with a DNA sequence encoding BMP-3, said DNA sequence being in relative association with an expression control sequence therefor and isolating BMP-3 from said culture medium.

16. A process according to Claim 15 wherein said DNA sequence comprises substantially the nucleotide sequence of Table IVA and IVB.

17. A cDNA sequence encoding BMP-1 comprising substantially the nucleotide sequence of Table VI or a sequence which hybridize thereto under stringent conditions and which upon expression codes for a protein exhibiting substantial properties of BMP-1.

18. A cDNA sequence encoding BMP-2 Class I comprising substantially the nucleotide sequence of Table VII or a sequence which hybridizes thereto under stringent conditions and which upon expression codes for a protein exhibiting substantial properties of BMP-2 Class I.

19. A cDNA sequence encoding BMP-2 Class II comprising

AMENDED CLAIMS
[received by the International Bureau on 8 December 1987 (08.12.87)
original claims 6, 8, 10, 12, 14, 16-20 amended;
new claims 21-23 added; other claims unchanged (13 pages)]

p.9 60

1. A pharmaceutical composition comprising a protein selected from the group consisting of:

(a) BMP-1;

(b) BMP-2 Class I;

(c) BMP-2 Class II;

(d) BMP-3; and

mixtures thereof, in a pharmaceutically acceptable vehicle.

2. A composition of Claim 1 wherein said protein is BMP-1.

3. A composition of Claim 1 wherein said protein is BMP-2 Class I.

4. A composition of Claim 1 wherein said protein is BMP-2 Class II.

5. A composition of Claim 1 wherein said protein is BMP-3.

6. The pharmaceutical composition of Claim 1 further comprising a matrix capable of delivering the composition to the site of the bone or cartilage defect and providing a structure for inducing bone or cartilage formation.

7. The composition of Claim 6 wherein said matrix comprises a material selected from the group consisting of hydroxyapatite, collagen, polylactic acid and tricalcium phosphate.

8. A method for inducing bone or cartilage formation in a patient in need of same comprising administering to said patient an effective amount of a composition of Claim 1-7.

```
                560                 575                 590
GAC GAG GAC AGC TAT ATT GTG TTC ACC TAT CGA CCT TGC GGG TGC TGC TCC TAC
Asp Glu Asp Ser Tyr Ile Val Phe Thr Tyr Arg Pro Cys Gly Cys Cys Ser Tyr

    605                 620                 635                 650
GTG GGT CGC CGC GGC GGG GGC CCC CAG GCC ATC TCC ATC GGC AAG AAC TGT GAC
Val Gly Arg Arg Gly Gly Gly Pro Gln Ala Ile Ser Ile Gly Lys Asn Cys Asp

            665                 680                 695
AAG TTC GGC ATT GTG GTC CAC GAG CTG GGC CAC GTC GTC GGC TTC TGG CAC GAA
Lys Phe Gly Ile Val Val His Glu Leu Gly His Val Val Gly Phe Trp His Glu

710                 725                 740                 755
CAC ACT CGG CCA GAC CGG GAC CGC CAC GTT TCC ATC GTT CGT GAG AAC ATC CAG
His Thr Arg Pro Asp Arg Asp Arg His Val Ser Ile Val Arg Glu Asn Ile Gln

        770                 785                 800                 815
CCA GGG CAG GAG TAT AAC TTC CTG AAG ATG GAG CCT CAG GAG GTG GAG TCC CTG
Pro Gly Gln Glu Tyr Asn Phe Leu Lys MET Glu Pro Gln Glu Val Glu Ser Leu

                830                 845                 860
GGG GAG ACC TAT GAC TTC GAC AGC ATC ATG CAT TAC GCT CGG AAC ACA TTC TCC
Gly Glu Thr Tyr Asp Phe Asp Ser Ile MET His Tyr Ala Arg Asn Thr Phe Ser

    875                 890                 905                 920
AGG GGC ATC TTC CTG GAT ACC ATT GTC CCC AAG TAT GAG GTG AAC GGG GTG AAA
Arg Gly Ile Phe Leu Asp Thr Ile Val Pro Lys Tyr Glu Val Asn Gly Val Lys

            935                 950                 965
CCT CCC ATT GGC CAA AGG ACA CGG CTC AGC AAG GGG GAC ATT GCC CAA GCC CGC
Pro Pro Ile Gly Gln Arg Thr Arg Leu Ser Lys Gly Asp Ile Ala Gln Ala Arg

980                 995                 1010                1025
AAG CTT TAC AAG TGC CCA GCC TGT GGA GAG ACC CTG CAA GAC AGC ACA GGC AAC
Lys Leu Tyr Lys Cys Pro Ala Cys Gly Glu Thr Leu Gln Asp Ser Thr Gly Asn

        1040                1055                1070                1085
TTC TCC TCC CCT GAA TAC CCC AAT GGC TAC TCT GCT CAC ATG CAC TGC GTG TGG
Phe Ser Ser Pro Glu Tyr Pro Asn Gly Tyr Ser Ala His MET His Cys Val Trp

                1100                1115                1130
CGC ATC TCT GTC ACA CCC GGG GAG AAG ATC ATC CTG AAC TTC ACG TCC CTG GAC
Arg Ile Ser Val Thr Pro Gly Glu Lys Ile Ile Leu Asn Phe Thr Ser Leu Asp

    1145                1160                1175                1190
CTG TAC CGC AGC CGC CTG TGC TGG TAC GAC TAT GTG GAG GTC CGA GAT GGC TTC
Leu Tyr Arg Ser Arg Leu Cys Trp Tyr Asp Tyr Val Glu Val Arg Asp Gly Phe

            1205                1220                1235
TGG AGG AAG GCG CCC CTC CGA GGC CGC TTC TGC GGG TCC AAA CTC CCT GAG CCT
Trp Arg Lys Ala Pro Leu Arg Gly Arg Phe Cys Gly Ser Lys Leu Pro Glu Pro
```

```
  1955                1970                1985                2000
GAG GGC AAT GAT GTG TGC AAG TAC GAC TTC GTG GAG GTG CGC AGT GGA CTC ACA
Glu Gly Asn Asp Val Cys Lys Tyr Asp Phe Val Glu Val Arg Ser Gly Leu Thr

          2015                2030                2045
GCT GAC TCC AAG CTG CAT GGC AAG TTC TGT GGT TCT GAG AAG CCC GAG GTC ATC
Ala Asp Ser Lys Leu His Gly Lys Phe Cys Gly Ser Glu Lys Pro Glu Val Ile

2060                2075                2090                2105
ACC TCC CAG TAC AAC AAC ATG CGC GTG GAG TTC AAG TCC GAC AAC ACC GTG TCC
Thr Ser Gln Tyr Asn Asn MET Arg Val Glu Phe Lys Ser Asp Asn Thr Val Ser

     2120                2135                2150                2165
AAA AAG GGC TTC AAG GCC CAC TTC TTC TCA GAA AAG AGG CCA GCT CTG CAG CCC
Lys Lys Gly Phe Lys Ala His Phe Phe Ser Glu Lys Arg Pro Ala Leu Gln Pro

               2180                2195                2210
CCT CGG GGA CGC CCC CAC CAG CTC AAA TTC CGA GTG CAG AAA AGA AAC CGG ACC
Pro Arg Gly Arg Pro His Gln Leu Lys Phe Arg Val Gln Lys Arg Asn Arg Thr

  2225      2235      2245      2255      2265      2275      2285
CCC CAG TGAGGCCTGC CAGGCCTCCC GGACCCCTTG TTACTCAGGA ACCTCACCTT GGACGGAATG
Pro Gln

     2295      2305      2315      2325      2335      2345      2355
GGATGGGGGC TTCGGTGCCC ACCAACCCCC CACCTCCACT CTGCCATTCC GGCCCACCTC CCTCTGGCCG

     2365      2375      2385      2395      2405      2415      2425
GACAGAACTG GTGCTCTCTT CTCCCCACTG TGCCCGTCCG CGGACCGGGG ACCCTTCCCC GTGCCCTACC

     2435      2445      2455      2465      2475      2485      2495
CCCTCCCATT TTGATGGTGT CTGTGACATT TCCTGTTGTG AAGTAAAAGA GGGACCCCTG CGTCCTGCCT

CTAGA
```

11. A process for producing BMP-2 Class I comprising culturing in a suitable culture medium a cell line transformed with a DNA sequence encoding BMP-2 Class I, said DNA sequence being in relative association with an expression control sequence therefor, and isolating BMP-2 Class I from said culture medium.

12. A process for according to Claim 11 wherein said DNA sequence comprises substantially the nucleotide sequence as

```
                    850                 865                 880
TTA GGA AAC AAT AGC AGT TTC CAT CAC CGA ATT AAT ATT TAT GAA ATC ATA AAA
Leu Gly Asn Asn Ser Ser Phe His His Arg Ile Asn Ile Tyr Glu Ile Ile Lys

    895                 910                 925                 940
CCT GCA ACA GCC AAC TCG AAA TTC CCC GTG ACC AGT CTT TTG GAC ACC AGG TTG
Pro Ala Thr Ala Asn Ser Lys Phe Pro Val Thr Ser Leu Leu Asp Thr Arg Leu

            955                 970                 985
GTG AAT CAG AAT GCA AGC AGG TGG GAA AGT TTT GAT GTC ACC CCC GCT GTG ATG
Val Asn Gln Asn Ala Ser Arg Trp Glu Ser Phe Asp Val Thr Pro Ala Val MET

1000                1015                1030                1045
CGG TGG ACT GCA CAG GGA CAC GCC AAC CAT GGA TTC GTG GTG GAA GTG GCC CAC
Arg Trp Thr Ala Gln Gly His Ala Asn His Gly Phe Val Val Glu Val Ala His

        1060                1075                1090                1105
TTG GAG GAG AAA CAA GGT GTC TCC AAG AGA CAT GTT AGG ATA AGC AGG TCT TTG
Leu Glu Glu Lys Gln Gly Val Ser Lys Arg His Val Arg Ile Ser Arg Ser Leu

    1105                1120                1135                1150
CAC CAA GAT GAA CAC AGC TGG TCA CAG ATA AGG CCA TTG CTA GTA ACT TTT GGC
His Gln Asp Glu His Ser Trp Ser Gln Ile Arg Pro Leu Leu Val Thr Phe Gly

    1165                1180                1195                1210
CAT GAT GGA AAA GGG CAT CCT CTC CAC AAA AGA GAA AAA CGT CAA GCC AAA CAC
His Asp Gly Lys Gly His Pro Leu His Lys Arg Glu Lys Arg Gln Ala Lys His

        1225                1240                1255
AAA CAG CGG AAA CGC CTT AAG TCC AGC TGT AAG AGA CAC CCT TTG TAC GTG GAC
Lys Gln Arg Lys Arg Leu Lys Ser Ser Cys Lys Arg His Pro Leu Tyr Val Asp

1270                1285                1300                1315
TTC AGT GAC GTG GGG TGG AAT GAC TGG ATT GTG GCT CCC CCG GGG TAT CAC GCC
Phe Ser Asp Val Gly Trp Asn Asp Trp Ile Val Ala Pro Pro Gly Tyr His Ala,

        1330                1345                1360                1375
TTT TAC TGC CAC GGA GAA TGC CCT TTT CCT CTG GCT GAT CAT CTG AAC TCC ACT
Phe Tyr Cys His Gly Glu Cys Pro Phe Pro Leu Ala Asp His Leu Asn Ser Thr

                1390                1405                1420
AAT CAT GCC ATT GTT CAG ACG TTG GTC AAC TCT GTT AAC TCT AAG ATT CCT AAG
Asn His Ala Ile Val Gln Thr Leu Val Asn Ser Val Asn Ser Lys Ile Pro Lys

    1435                1450                1465                1480
GCA TGC TGT GTC CCG ACA GAA CTC AGT GCT ATC TCG ATG CTG TAC CTT GAC GAG
Ala Cys Cys Val Pro Thr Glu Leu Ser Ala Ile Ser MET Leu Tyr Leu Asp Glu
```

```
522                 537                 552                 567
GGC GAC GCG GGA GGA CGC CGC TCA GGG CAG AGC CAT GAG CTC CTG CGG GAC TTC
Gly His Ala Gly Gly Arg Arg Ser Gly Gln Ser His Glu Leu Leu Arg Asp Phe

    582                 597                 612                 627
GAG GCG ACA CTT CTG CAG ATG TTT GGG CTG CGG CGC CGC CCG CAG CCT AGC AAG
Glu Ala Thr Leu Leu Gln MET Phe Gly Leu Arg Arg Arg Pro Gln Pro Ser Lys

                642                 657                 672
AGT GCC GTC ATT CCG GAC TAC ATG CGG GAT CTT TAC CGG CTT CAG TCT GGG GAG
Ser Ala Val Ile Pro Asp Tyr MET Arg Asp Leu Tyr Arg Leu Gln Ser Gly Glu

    687                 702                 717                 732
GAG GAG GAA GAG CAG ATC CAC AGC ACT GGT CTT GAG TAT CCT GAG CGC CCG GCC
Glu Glu Glu Glu Gln Ile His Ser Thr Gly Leu Glu Tyr Pro Glu Arg Pro Ala

            747                 762                 777
AGC CGG GCC AAC ACC GTG AGG AGC TTC CAC CAC GAA GAA CAT CTG GAG AAC ATC
Ser Arg Ala Asn Thr Val Arg Ser Phe His His Glu Glu His Leu Glu Asn Ile

792                 807                 822                 837
CCA GGG ACC AGT GAA AAC TCT GCT TTT CGT TTC CTC TTT AAC CTC AGC AGC ATC
Pro Gly Thr Ser Glu Asn Ser Ala Phe Arg Phe Leu Phe Asn Leu Ser Ser Ile

    852                 867                 882                 897
CCT GAG AAC GAG GCG ATC TCC TCT GCA GAG CTT CGG CTC TTC CGG GAG CAG GTG
Pro Glu Asn Glu Ala Ile Ser Ser Ala Glu Leu Arg Leu Phe Arg Glu Gln Val

            912                 927                 942
GAC CAG GGC CCT GAT TGG GAA AGG GGC TTC CAC CGT ATA AAC ATT TAT GAG GTT
Asp Gln Gly Pro Asp Trp Glu Arg Gly Phe His Arg Ile Asn Ile Tyr Glu Val

957                 972                 987                1002
ATG AAG CCC CCA GCA GAA GTG GTG CCT GGG CAC CTC ATC ACA CGA CTA CTG GAC
MET Lys Pro Pro Ala Glu Val Val Pro Gly His Leu Ile Thr Arg Leu Leu Asp

        1017                1032                1047
ACG AGA CTG GTC CAC CAC AAT GTG ACA CGG TGG GAA ACT TTT GAT GTG AGC CCT
Thr Arg Leu Val His His Asn Val Thr Arg Trp Glu Thr Phe Asp Val Ser Pro

1062                1077                1092                1107
GCG GTC CTT CGC TGG ACC CGG GAG AAG CAG CCA AAC TAT GGG CTA GCC ATT GAG
Ala Val Leu Arg Trp Thr Arg Glu Lys Gln Pro Asn Tyr Gly Leu Ala Ile Glu

       1122                1137                1152                1167
GTG ACT CAC CTC CAT CAG ACT CGG ACC CAC CAG GGC CAG CAT GTC AGG ATT AGC
Val Thr His Leu His Gln Thr Arg Thr His Gln Gly Gln His Val Arg Ile Ser

                1182                1197                1212
CGA TCG TTA CCT CAA GGG AGT GGG AAT TGG GCC CAG CTC CGG CCC CTC CTG GTC
Arg Ser Leu Pro Gln Gly Ser Gly Asn Trp Ala Gln Leu Arg Pro Leu Leu Val
```

association with an expression control sequence therefor and isolating BMP-3 from said culture medium.

16. A process according to Claim 15 wherein said DNA sequence comprises substantially the nucleotide sequence as follows:

```
       383        393        403        413                  428
GAGGAGGAAG CGGTCTACGG GGGTCCTTCT GCCTCTGCAG  AAC AAT GAG CTT CCT GGG GCA
                                             Asn Asn Glu Leu Pro Gly Ala

       443                   458                  473                  488
GAA TAT CAG TAC AAG GAG GAT GAA GTA TGG GAG GAG AGG AAG CCT TAC AAG ACT
Glu Tyr Gln Tyr Lys Glu Asp Glu Val Trp Glu Glu Arg Lys Pro Tyr Lys Thr

                  503                  518                  533
CTT CAG ACT CAG CCC CCT GAT AAG AGT AAG AAC AAA AAG AAA CAG AGG AAG GGA
Leu Gln Thr Gln Pro Pro Asp Lys Ser Lys Asn Lys Lys Lys Gln Arg Lys Gly

    548                  563                  578                  593
CCT CAG CAG AAG AGT CAG ACG CTC CAG TTT GAT GAA CAG ACC CTG AAG AAG GCA
Pro Gln Gln Lys Ser Gln Thr Leu Gln Phe Asp Glu Gln Thr Leu Lys Lys Ala

            608                  623                  638
AGA AGA AAG CAA TGG ATT GAA CCC CGG AAT TGT GCC AGA CGG TAC CTT AAA GTG
Arg Arg Lys Gln Trp Ile Glu Pro Arg Asn Cys Ala Arg Arg Tyr Leu Lys Val

653                  668                  683                  698
GAC TTC GCA GAT ATT GGC TGG AGC GAA TGG ATT ATT TCC CCC AAG TCC TTC GAT
Asp Phe Ala Asp Ile Gly Trp Ser Glu Trp Ile Ile Ser Pro Lys Ser Phe Asp

       713                  728                  743                  756        766
GCC TAT TAC TGC TCC GGA GCG TGC CAG TTC CCC ATG CCA AAG GTAGCCATTG TTTTTTGTCC
Ala Tyr Tyr Cys Ser Gly Ala Cys Gln Phe Pro MET Pro Lys

       776        786
TGTCCTTCCC ATTTCCATAG ; and

       284        294        304                  319
CTAACCTGTG TTCTCCCTTT TCGTTCCTAG  TCT TTG AAG CCA TCA AAT CAC GCT ACC
                                  Ser Leu Lys Pro Ser Asn His Ala Thr

334                  349                  364                  379
ATC CAG AGT ATA GTG AGA GCT GTG GGG GTC GTC CCT GGA ATC CCC GAG CCT TGC
Ile Gln Ser Ile Val Arg Ala Val Gly Val Val Pro Gly Ile Pro Glu Pro Cys

       394                  409                  424                  439
TGT GTG CCA GAA AAG ATG TCC TCA CTC AGC ATC TTA TTC TTT GAT GAA AAC AAG
Cys Val Pro Glu Lys MET Ser Ser Leu Ser Ile Leu Phe Phe Asp Glu Asn Lys
```

This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☒ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.
As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.